Journal of Business Data Science Research

A Machine Learning Algorithm for Money Laundering Detection in Bank Melli Iran

Mehdi Shakeri Behbahani[a], Naser Khani[b,*]

[a] Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
[b] Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran.

ABSTRACT

In the study, different feature selection methods were initially studied to prevent and detect money laundering, and then a new method was developed and used in three stages for the selection of features effective in detecting money laundering using a cellular learning automata-based algorithm. In the first stage, the patterns were extracted using paired features through a complete graph. In the second stage, the extracted patterns were trained and classified on the basis of the impact rate of features using the cellular learning automata (CLA). Finally, in the third stage, the optimized feature was selected based on the impact rate of features. Selection of effective features using the proposed method improved the accuracy of data classification to detect money laundering. The Bank Melli Iran data set was utilized by entering into MATLAB to evaluate the proposed method and compare it with other methods. The results showed that the accuracy rate of classification in the proposed CLA method to detect money laundering was 94.19% and its runtime was 263.32 seconds. The proposed method was observed to have higher classification accuracy in detecting money laundering, as compared to the listed methods.

KEWWORDS: Feature selection, cellular learning automata, machine learning, money laundering detection.

ORGINAL ARTICLE

Received 19 July 2020
Accepted 24 February 2021

CITATION

Shakeri Behbahani, M., Khani, N. (2021). A Machine Learning Algorithm for Money Laundering Detection in Bank Melli Iran. *Journal of Business Data Science Research*, 1(1), 5-13.

1. Introduction

Money laundering is defined as the legitimization of money that is obtained illegally and illegitimately. The advent of electronic banking is a promising phenomenon with many benefits and advantages compared to older methods of banking since it can dramatically save people's time, energy, and capital. This newly emerged phenomenon is also beneficial to offenders and so amazing to money launderers who call it their heaven. Criminals and pioneers in money-laundering have made use of electronic banking in many ways to legitimize their dirty money. In electronic money laundering, instead of conventional troublesome money laundering methods, money launderers make use of a group of techniques and cyberspace, such as quick transfer of money from one country to other countries, to clean the sources of their proceeds. Unfortunately, in Iran, despite weaknesses in cyberspace and electronic banking systems and the possibility of their misuse for money laundering purposes, no essential measures have yet been taken to prevent electronic money laundering and the existing anti-money laundering regulations are not sufficiently deterrent and guaranteed to be enforced. Therefore, serious and extensive legislation and measures, as well as international cooperation, are required in the field to prevent and control this type of money laundering. This study proposes a smart solution for data mining, as well as detecting and preventing money laundering using the cellular learning automata (CLA). To do so, a cellular learning automata-based algorithm was proposed to detect money laundering in Bank Melli Iran. Accordingly, the initial data was entered into the CLA and after sorting, the effective pattern was extracted in the complete graph-based CLA. Subsequently, to improve the feature selection features not affecting the CLA classification (with an impact rate of less than the threshold 0.4) were excluded.

* Corresponding author: naserkhani@phu.iaun.ac.ir
• http://www.jbdsr.com/article_680444.html

2. Theoretical foundation

The term ‘money laundering’ is a historical phenomenon that has newly entered into economic concepts and has been subject to penalties. Two thousand years BC, merchants hid their assets from rulers who seized their wealth and destroyed them. Chinese merchants, in addition to hiding their assets from the government, also transferred them. Although the term was coined in the twentieth century, its rules were regulated many years ago. Sterling Shigrov in one of his books has discussed China's history. He explains how the abuse of rulers from merchants and others had caused them to look for ways to hide their wealth so that they can transfer them from one place to another without being detected. Taken all these into account, the act of money laundering goes back to 4000 years BC.

Ansari Pirsaraie and Shahbahrami (2014) in an article discussed the necessity of using money laundering detection systems in electronic banking. In today's world, it is possible to use information technology in all aspects of human life. One of these areas is the e-banking industry that has enabled banking operations to be conducted quickly, accurately, and without time and place limitation at any moment of the day. However, the use of this technology entails some risks, including the act of money laundering.

Money launderers attempt to make the most use of virtual and cyberspace in e-banking in line with their criminal activities, as obscure money transfers without intermediaries could lead to a safety margin for them. With advances in electronic banking, money laundering methods have gradually become more sophisticated and thus, are impossible to detect without the use of money laundering detection systems due to the massive amount of data.

Afrash and Khodamoradi (2014) in their article explored the use of data mining for anti-money laundering practices and fraud detection. Over the past few decades, the use of automated analysis tools for data mining in the field of fraud detection has been increasingly regarded as an important factor in the growth of data volumes. This project describes a specific method to learn how to detect existing patterns in relational data using graph display. Our research in the future will lead to the creation of a web-based application that allows users to search suspicious patterns through graphical queries in relational data.

Sharifi (2013) in an article investigated money laundering and strategies to deal with it. Money laundering is an invisible crime with no boundaries. Technology has led to the development of new methods to hide the source of proceeds resulting from criminal acts. ATM machines have made it extremely difficult, if not impossible, to track illicit money. Therefore, it is necessary to perform coordinated actions against money laundering through controlling and supervising banks and financial institutions.

Naheem (2016) in an article examined anomaly detection methods in the financial sphere. Anomaly detection is the important task of data analysis. One of the applications is to detect interesting patterns emerging in the process of data anomalies. Anomaly detection is an important tool for detecting abnormalities in many different areas, including financial fraud detection, intrusion into computer networks, human behavioral analysis, gene expression analysis, and many more. Recently in the financial sector, greater attention has been paid to detect fraudulent activities. Huge work has been carried out in anomaly detection-based clustering without being supervised in the financial sphere. This study aims at conducting an in-depth review of various methods of clustering based on the detection of anomalies from different perspectives. Yet in another article, Derzhesky et al. (2012) investigated money laundering detection systems.

Criminal analysis of information collected from various sources is a complex process, not in quantitative characteristics such as invoice issuance or bank account transactions, but in qualitative characteristics such as eyewitness testimony. Given the widespread nature of this information, operational activities or examinations could be much improved through using support tools and specific techniques.

Kumar and Arora (2016) in an article examined a combined method with the use of maximum entropy and Bayesian learning to detect criminals engaged in the financial industry. The use of credit cards because of the economic boom in the past few years has been tremendously increased; this has also resulted in an increase in cases of credit card fraud. The leading banks and software development companies are doing serious measures to deal with the situation.

They proposed in their study a framework for the detection of fraudulent use of credit cards where fraud is detected through maximum entropy due to the erratic behavior of customers in various credit card trades. Considering their parallel nature (which allows special-purposed hardware to be implemented), CA is also called to be a breakthrough in paradigms such as Smart Dust (Ilyas & Mahgoub, 2018; Warneke et al., 2001), Utility Fog (Dastjerdi & Buyya, 2016; Hall, 1996), Microelectromechanical Systems (MEMS or “motes”) (Judy, 2001), or Swarm Intelligence and Robotics (Del Ser et al., 2019), due to their capability to be computationally complete. Microscopic sensors are set to revolutionize a range of sectors, such as space missions (Niccolai et al., 2019). The sensing, control, and learning algorithms that need to be embarked on such miniaturized devices can currently only be run on relatively heavy hardware, and need to be refined. However, nature has shown us that this is possible in the brains of insects with only a few hundred neurons. CA can also be decisive in other paradigms such as Nanotechnology, Ubiquitous Computing (López-de-Ipina et al., 2017), and Quantum Computation introduced by Feynman (1986), particularly the Quantum CA (Adamatzky, 2018). Due to the miniaturization hurdles of these devices, CA for stream learning may become of interest when there is no enough capacity to store huge volumes of data, and the computational capacity is very limited. After all, it is therefore not surprising that CA has received particular attention since the early days of CA investigation, and that CA for stream learning allows us to move in the correct direction in these scenarios, which are not far off the near future (Jafferis et al., 2019).

With the capability of evolving in a simple structure, CA has been proven to produce different complex phenomena in a natural system in robust and efficient ways (Khomami et al., 2018). As an extension of CLA, dynamic cellular learning automaton (DCLA) is proposed for more extensive applications (Saghiri & Meybodi, 2017).

3. The proposed method

The research method is in fact the design of the research for achieving the research goals. In this study, using the CLA, data extracted from information and performance statement of Bank Melli Iran were data-mined, and evaluated using statistical analysis of the model presented in this study. Finally, after data mining with the use of the CLA, an artificial neural network approach was used to assess the proposed method and present the prediction model.

Data mining using the CLA is generally divided into three stages:
1. Extracting patterns using a complete graph
2. Extracting useful patterns in the complete graph-based CLA
3. Selecting features based on the impact rate of patterns extracted in the CLA
The three stages are as follows:

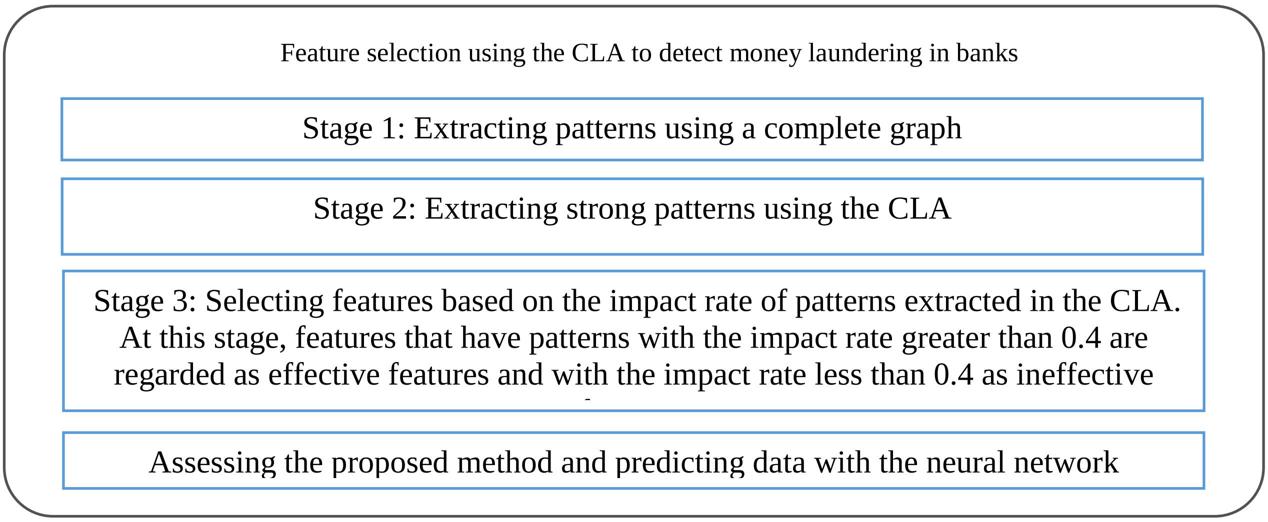

Figure 1. The process of the proposed CLA method

3.1. Extracting patterns using a complete graph

Complete graph is a graph where each node is connected to another and a graph is assigned to a class of datasets. A node represents a feature and an edge shows a pair of related features. An edge pattern and weight reflect the impact rate of patterns.

- Implementation of a pair of related features

At this stage, a list of impact rate and information was recorded for data possessing an effective and specific pattern using the following format:

$$\langle \text{Data\#} \rangle \quad \langle \text{atr\#} \rangle.\langle \text{atrValue} \rangle \rightarrow \langle \text{atr\#} \rangle.\langle \text{atrValue} \rangle \quad \langle \text{EffectiveRate} \rangle \quad \langle \text{Class} \rangle$$

<Class> < impact rate > <attribute value>. <Attribute j> <attribute value>. <Attribute i> <data (records)>

- A sample of patterns extracted using a complete graph:

In Table 1 for the datasets containing 5 rows and 4 features, the data values ranges between 0 to 3, with two categories or classes (0 and 1).

In Table 1 for Class 1, the features # 2 and # 1 of data in the first raw are similar to the features # 1 and # 2 of the fourth raw (number of bold letters in Table 1). Therefore, the information is recorded as follows:

$$1,4 \quad 1.1 \rightarrow 2.0$$

The impact rate of alpha in the above pattern is 2 because it has two sets of related features with values (1, 4) and the relevant information are directly copied in the table indicating the pair of related features so that the impact rates of its class number are 2 and 1, respectively.

If there is no similarity among the pairs of related features of two or more data, it can be recorded as the only sample of information. For example, the features 1 and 4 of the data in the first raw are not similar to any other information; hence, this pattern is recorded as follows:

$$1 \quad 1.1 \rightarrow 4.2$$

Table 1. Sample of pilot dataset

No.	Feature 1	Feature 2	Feature 3	Feature 4	Class
1	1	0	1	2	1
2	2	0	2	3	0
3	2	0	1	3	0
4	1	0	1	1	1
5	2	0	1	1	1

In the above pattern, the impact and class rates are 1 and 1, respectively, recorded as a pair of related features (Table 2). Since the dataset has four features, there should be 4 pairs of attribute values; hence, each data should appear 4 times in the extraction table. For example, if there is N features for the data row 1 in Table 1, the number of rows with similar data to be included as a pair of related features (table 2) is shown as the following equation.

$${}^{n}C_{2} = \frac{n!}{(n-2)!2!} \tag{1}$$

Table 2. Pairs of related features

No.	Pairs of related features	Impact Rate	Class
1, 4	$1.1 \rightarrow 2.0$	2	1
1, 4	$1.1 \rightarrow 3.1$	2	1
1	$1.1 \rightarrow 4.2$	1	1
1	$2.0 \rightarrow 4.2$	1	1
1	$3.1 \rightarrow 4.2$	1	1
1, 4, 5	$2.0 \rightarrow 3.1$	3	1
2, 3	$1.2 \rightarrow 2.0$	2	0
2, 3	$1.2 \rightarrow 4.3$	2	0
2, 3	$2.0 \rightarrow 4.3$	2	0
2	$1.2 \rightarrow 3.2$	1	0
2	$2.0 \rightarrow 3.2$	1	0
2	$3.2 \rightarrow 4.3$	1	0
3	$1.2 \rightarrow 3.1$	1	0
3	$2.0 \rightarrow 3.1$	1	0
3	$3.1 \rightarrow 4.3$	1	0

4	$1.1 \rightarrow 4.1$	1	1
4	$2.0 \rightarrow 4.1$	1	1
4	$3.1 \rightarrow 4.1$	1	1
5	$1.2 \rightarrow 2.0$	1	1
5	$1.2 \rightarrow 3.1$	1	1
5	$1.2 \rightarrow 4.1$	1	1
5	$2 \rightarrow 4.1$	1	1
5	$3.1 \rightarrow 4.1$	1	1

To extract Table 2 from Table 1, the following steps were taken:
For each class of data rows of datasets, a complete graph of features is applicable.
Nodes represent features and edge shows the dependence of pairs of related features and node status is indicative of the value of features. For example, for features 1 with data value 2, there is a node for feature 1 (A + R # 1) in two modes (1 and 2).

Edges are created by ATR # 1.1 and ATR # 2.0 creates a graph.
Repeated edges indicate the impact rate. For example, if an edge is created twice, its impact rate is 2 times.
The same is separately done for all classes and the first and second columns of Table 2 are achieved after running the algorithm.

3.2. Effective pattern extraction using the CLA

At this stage, the CLA is used to extract effective patterns in each class. The CLA can combine two patterns and extract their impact rate. A cell of the CLA represents the data row number and a group of data having same pattern. Each cell of the CLA has a LA recording the adjacency rate for each cell. In each LA, if i$\boldsymbol{a}$ is a set of row numbers, β (n) = 1 will be a reward, and if "adjacency rate> compliance value, (n) = 0β and the penalty pattern will be received.
The compliance rate is determined based on 50% compliance between two neighboring groups. Adjacency rates for the occurrence of similar neighbors are similar. In any implementation model, effective dependencies/ outcomes of the first stage (Pi) are converted into an input to the second stage.

Adjacency rate refers to the occurrence of each row in Table 2; however, the compliance rate in the CLA process is obtained according to $\boldsymbol{a}$ estimated in the previous step. Adjacency rate of each cell is obtained from dividing the number of comparative rows for each pair of related features by the total number of rows for each pair of features. If the adjacency rate is greater than a specified percentage, there will be β (n) = 1 and is regarded as a reward; otherwise, it is penalty and there will be β (n) = 0.

In this phase, the reward rate is equal with the penalty rate ($\boldsymbol{a}$ and b are equal to 0.5) and given that the rate of rewards and penalties are equal, all patterns in the LRP learning model have the following operation probability value:

$$\forall i,\ i <= n \quad \rho_i = \frac{1}{n} \tag{2}$$

Where, *n* is the number of pairs of related features.
The output of the second stage is a pattern of corresponding dataset in which the equation (3) was used to achieve the reward pattern; otherwise, the penalty is recorded using equation (4).

$$P_i(n+1) = P_i(n) + \alpha[1 - P_i(n)] \tag{3}$$

$$P_i(n+1) = (1-b)P_i(n) \tag{4}$$

Given the row number of effective patterns, whose Pi value is greater than 0.5 during the CLA process and are optimal, the pattern, its features and impact rate are the content of relatedness in the pattern (Table 1). Accordingly, the amount of reinforcement can be calculated.

A sample of effective pattern extraction:

After extracting effective patterns in Table 1, 1 and 4 are adjacent and their compatibility rate with 1, 4 and 5 is above 50%. Thus, the adjacencies can be combined according to this rule.

After reading the first row of Table 2, 1 and 4 are adjacent for the first time in the CLA. Then, if rewards are received from other neighbors, there will be a compliance rate greater than compatibility rate. This process is done for each model having greater number of members (Esmaeil Pour et al., 2012).

(Total number of rows for each pair of related features) / (number of matching columns for each pair of related features) = Adjacency Rate (5)

If the probability is above 0.5, this means that it is an effective adjacency which is reinforced and all pairs of its related features are combined. Table 3 (a) and Table 3 (b) will be achieved after combining pairs of related features.

Table 3. (a) Developing an effective model for class 1, (b) Developing an effective model for class 0

(a)

Row No.	Dependency of the pair of features	Impact rate	Amount of reward (reinforcement)
1,4	1.1→2.0 , 1.1→3.1 , 2.0→3.1	2	3*2 = 6
1 , 4 , 5	1.1→2.0 , 1.1→3.1 , 2.0→3.1 And 2.0 → 3.1	2 1	3*2 = 6 → 6+1 = 7 1 * 1 = 1

(b)

Row No.	Dependency of the pair of features	Impact rate	Amount of reward (reinforcement)
2 , 3	1.2 → 2.0 , 1.2 → 4.3 , 2.0 → 4.3	2	3 * 2 = 6

If there is "adjacency rate>compliance rate" for a new adjacency, then it is rewarded and they will be classified in the same group with previous data. In other words, a more effective pattern is created. Otherwise, the penalty is received and this leads to a decreased relationship between previous cells. At the time, the same occurrence is transferred to another group and the relationship between previous neighbors reduces.

3.3. Feature selection based on the impact rate of extracted patterns using the CLA

Given that the probability vector showing the relationship of the classes *a* and *b* is driven from Table 3, each pair of related features has an impact value on the probability vector.

The most important feature selection criteria are as follows:

1. Only features existing in the class a or class b probability vector are selected and other features not existing in these two vectors are removed at this stage (features that have no role in the classification of classes are excluded).

2. At the second stage, given that each row of the dependency table has a certain impact rate, features with the impact rate less than 40% are removed.

Ultimately, final features obtained can be compared with other methods of feature selection such as the neural network, support vector machine and decision tree through using classifiers.

4. Results and Tests

In this section, we analyze the data using the proposed method and compare it with other methods. This study was to predict and discover the customers and accounts which are suspected of money laundering using the proposed method. After evaluating and comparing the results of the proposed method with the other methods in the MATLAB software, the classification was improved in terms of accuracy and mean square error of categories.

The data used in the study were extracted from time-series data of Bank Melli Iran, the data of the Statistical Center of Iran and the Melli Bank's performance reports. One of the major issues is to adjust the data time periods. Parameters of the above data are displayed below.

4.1. Preliminary values of dataset and the proposed method

To review the proposed method, the findings should be evaluated for the Melli Bank's dataset. The following table shows the number of data, features and classes as well as the number of features obtained from the proposed method.

Table 4. Preliminary data and number of features

<table>
<tr><th rowspan="2">Datasets</th><th colspan="3">Preliminary number</th><th colspan="11">Number of features</th></tr>
<tr><th>data</th><th>features</th><th>class</th><th colspan="11">CLA</th></tr>
<tr><td rowspan="2">Melli bank dataset</td><td rowspan="2">500</td><td rowspan="2">11</td><td rowspan="2">2</td><td colspan="11">9</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td></tr>
</table>

For the evaluated data shown in Table 4, the preliminary number of data, features and classes are also represented. Using the proposed method, the number of features decreased from 11 to 9.

4.2. Accuracy and mean square error of proposed method

To evaluate data mining using the CLA, the findings obtained from the proposed method were classified using the neural network. Then, the accuracy, error and runtime were estimated for the proposed method. Comparing the findings with the preliminary data revealed the improvement rate for the proposed method.

Table 5. Accuracy and mean square error of proposed method

Datasets	Results of the proposed method			
	Accuracy		Mean Square Error	
	Preliminary data	CLA	Preliminary data	CLA
Melli bank dataset	86.79	94.19	0.2984	0.1078

Three-layer Back Propagation neural network with 10 neurons was used to classify the data and 70 percent of the data was for training and 30% of the data was randomly selected for testing. As shown in Table 4-5, the proposed method, in addition to reducing the data dimensions, improved the classification accuracy for the evaluated dataset.

4.3. Accuracy of the proposed method compared to other methods

To evaluate the efficacy and accuracy of data mining using the CLA method, the accuracy of the proposed method is compared to the accuracy of other methods through using feature selection methods in Table 8.4 (Mousavirad & Ebrahimpour-Komleh, 2014).

Table 6. Comparison of the proposal method with other methods

Datasets	FFS (Koller & Sahami, 1996)	BFS (Koller & Sahami, 1996)	GA-FS (Yang & Honavar, 1998)	PSO-FS (Wang et al., 2007)	COA-FS (Mousavirad & Ebrahimpour-Komleh, 2014)	Proposed CLA Method
Melli bank dataset	74.32	75.65	80.95	82.45	83.18	94.19

In Table 6, the classification accuracy of the Melli Bank's dataset for the proposed method in comparison to the COA (Cuckoo Optimization Algorithm), PSO-FS (Particle Swarm Based Feature Selection), GA-FS (Genetic Based Feature Selection), BFS (Backward Feature Selection) and FFS (Forward Feature Selection) algorithms is presented.

The classification accuracy of the proposed method for the Melli Bank's high-dimensional datasets was 94.19%, which is more efficient compared to other methods. The proposed CLA method is more preferred for high dimensional data having more than two categories and its runtime is 263.32 seconds.

5. Conclusions and future work

Regarding the classification accuracy of the proposed CLA in comparison with the preliminary data and COA (Cuckoo Optimization Algorithm), PSO-FS (Particle Swarm Based Feature Selection), GA-FS (Genetic Based Feature Selection), BFS (Backward Feature Selection) and FFS (Forward Feature Selection) methods, the findings revealed that the proposed model categorized Bank Melli Iran data with greater accuracy and less error and acted better than other methods in detecting money laundering. The CLA method is more preferred for high dimensional data having more than two categories in terms of its runtime.

Although the COA method had a relatively greater accuracy for low-dimensional data compared to the proposed method, the COA method is only preferred for data having binary categories and it is not appropriate for large number of categories, e.g., five categories, since it requires less accuracy and more implementation time.

References

Adamatzky, A. (2018). *Cellular Automata: A Volume in the Encyclopedia of Complexity and Systems Science*. Springer.

Afrash, M., & Khodamoradi, M. (2014). *Use of data mining for anti-money laundering and fraud detection* National Conference on Engineering Sciences, New Ideas, Tonekabon.

Ansari Pirsaraie, Z., & Shahbahrami, A. (2014). The need to use money laundering detection systems in electronic banking. *Trend (Economic Research Trend)*, *21*(68), 179-211.

Dastjerdi, A. V., & Buyya, R. (2016). Fog computing: Helping the Internet of Things realize its potential. *Computer*, *49*(8), 112-116.

Del Ser, J., Osaba, E., Molina, D., Yang, X.-S., Salcedo-Sanz, S., Camacho, D., Das, S., Suganthan, P. N., Coello, C. A. C., & Herrera, F. (2019). Bio-inspired computation: Where we stand and what's next. *Swarm and Evolutionary Computation*, *48*, 220-250.

Feynman, R. P. (1986). Quantum mechanical computers. *Foundations of physics*, *16*(6), 507-531.

Hall, J. S. (1996). Utility fog: The stuff that dreams are made of. *Nanotechnology*, 161-184.

Ilyas, M., & Mahgoub, I. (2018). *Smart Dust: Sensor network applications, architecture and design*. CRC press.

Jafferis, N. T., Helbling, E. F., Karpelson, M., & Wood, R. J. (2019). Untethered flight of an insect-sized flapping-wing microscale aerial vehicle. *Nature*, *570*(7762), 491-495.

Judy, J. W. (2001). Microelectromechanical systems (MEMS): fabrication, design and applications. *Smart materials and Structures*, *10*(6), 1115.

Khomami, M. M. D., Rezvanian, A., & Meybodi, M. R. (2018). A new cellular learning automata-based algorithm for community detection in complex social networks. *Journal of computational science*, *24*, 413-426.

Koller, D., & Sahami, M. (1996). *Toward optimal feature selection*.

Kumar, D., & Arora, S. (2016). A hybrid approach using maximum entropy and bayesian learning for detecting delinquency in financial industry. *International Journal of Knowledge-Based Organizations (IJKBO)*, *6*(1), 60-73.

Mousavirad, S., & Ebrahimpour-Komleh, H. (2014). Wrapper feature selection using discrete cuckoo optimization algorithm. *International Journal of Mechatronics, Electrical, and Computer Engineering*, *4*(11), 709-721.

Naheem, M. A. (2016). Money laundering: A primer for banking staff. *International Journal of Disclosure and Governance*, *13*(2), 135-156.

Niccolai, L., Bassetto, M., Quarta, A. A., & Mengali, G. (2019). A review of Smart Dust architecture, dynamics, and mission applications. *Progress in Aerospace Sciences*, *106*, 1-14.

Saghiri, A. M., & Meybodi, M. R. (2017). A closed asynchronous dynamic model of cellular learning automata and its application to peer-to-peer networks. *Genetic Programming and Evolvable Machines*, *18*(3), 313-349.

Sharifi, S. (2013). *Money Laundering and Countermeasures* 2nd International Conference on Management, Entrepreneurship and Economic Development, Qom, Iran.

Wang, X., Yang, J., Teng, X., Xia, W., & Jensen, R. (2007). Feature selection based on rough sets and particle swarm optimization. *Pattern recognition letters*, *28*(4), 459-471.

Warneke, B., Last, M., Liebowitz, B., & Pister, K. S. (2001). Smart dust: Communicating with a cubic-millimeter computer. *Computer*, *34*(1), 44-51.

Yang, J., & Honavar, V. (1998). Feature subset selection using a genetic algorithm. In *Feature extraction, construction and selection* (pp. 117-136). Springer.